Big Data Analytics and Edge Computing

Arnav Arnav
Indiana University, Bloomington
Bloomington, Indiana, USA
aarnav@iu.edu

ABSTRACT

With the exponential increase in the number of connected IoT devices, the data generated by these devices has grown enormously. Sending this data to a centralized server or cloud results in enormous network traffic and may lead to failures and increased latency. A solution for this problem is to do some processing on the devices closer to the network edge, enabling responsive and real-time analytics. There have been various developments in the field of edge computing some of which are described here.

KEYWORDS

i523, hid201, Edge Computing, Big Data Analytics

1 INTRODUCTION

Internet of things is rapidly gaining importance and Evans Data Corporationfis Global Developer Population and Demographics Study reports 6.2 million developers working in the IoT domain[11]. With the rapid increase in the acceptance of Internet of Things (IoT) devices across various fields across the world, ranging from industrial sensors to lifestyle and sports products, and the consequent increase in the data generated by such devices, there is a pressing demand for devices and processes that can analyze this data and provide responsive analytics[13]. Traditionally, IoT applications follow one of the two approaches - cloud-centric approach, where the sensing devices send data to the cloud where the analytics are performed or device-centric approach, where stand-alone devices running proprietary code perform analytics locally[13]. Networks are largely centralized with organizations soring all data, which may not be directly beneficial to them, in their data centers, and data flowing from the edge to the cloud on each operation[2].

With an increase in the number of connected devices, it gets increasingly difficult to perform all analytics on a server in a traditional manner. Thus, edge computing involves pushing a part of this computation closer to the end user of the device, or closer to the network edge[14][2]. This helps reduce the cost incurred in communicating large amounts of data over the network, ensures some level of availability even when the connection to the cloud is broken and reduces the cost of computation and storing data on the cloud[13][2].

2 HOW EDGE COMPUTING WORKS

Edge computing emerged with the development of content delivery networks (CDNs) by Akamai which use nodes close to the user to prefetch web content and accelerate web throughput. Edge computing extends this concept with the help of cloud infrastructure to run arbitrary task-specific code at nodes close to the edge, typically known as cloudlets. These cloudlets usually run on a virtual machine or a light-weight container for ease of isolation and resource management[12].

Proximity to the edge of the network ensures various benefits. It helps to provide highly responsive applications, by using a more powerful computing resource near the edge and minimizing end-to-end latency, which is essential in time-critical applications like virtual reality which require a latency of less than 16ms for the images to appear stable[3][12]. Proximity also increases scalability with the help of edge analytics where cloudlets perform the first level of analytics on the sensor data and only send processed data and metadata to the cloud to reduce bandwidth usage as the number of connected devices increases[12]. Decentralization of data can also provide the owners of data more control over the privacy of their data, and provide ways to safely communicate this data between various entities[2][5].

In industrial applications like aviation where a large amount of data is generated on each flight[12], analyzing this data in a centralized manner becomes impractical. In such cases, fog computing is more useful which adds different elements at various levels of hierarchy between the edge and the cloud[7]. In industrial environments, there are a lot of different systems running new as well as legacy applications which may be proprietary and integrating these applications to provide end-to-end IoT solutions is still a challenge. Linux Foundation's EdgeX platform provides a way to simplify and standardize edge computing architectures and is gaining importance as an industrial IoT solution[7].

3 SOME EXAMPLES

Simmhan describes an application that was built using Apache-NiFi, a lightweight dataflow execution engine used for vehicle classification from video streams using a Tensorflow deep neural network encapsulated within a NiFi dataflow executing across multiple raspberry pis. This allows video streams to be analyzed locally and also provides the flexibility to use cloud infrastructure for computation when edge devices are constrained[13].

Yang Zhao et al proposed an occupancy and activity monitoring application with doppler sensing and edge analytics. The application uses low-cost motion sensing and embedded signal processing, detection and machine learning to detect activity in real time, even when multiple people are present in a room. The developers provide a web portal to help ease monitoring activity from a remote location[16].

Analysing video feeds on a large scale in real time is a challenging task. Each of the videos may be very large and a large amount of bandwidth is needed to stream the video feed to a central location which is not feasible especially if the cameras are connected wirelessly. In addition to this, the entire video may not be useful and most parts of it may be discarded depending on the application.

Furthermore, these applications need to provide results with low latency as important decisions often need to be made based on the output in case of surveillance applications[3]. Thus compute abilities available on cameras can be utilized to provide real-time video analysis, processing the video at the camera and only communicating interesting bits to the cloud[12].

A real-time video processing solution is proposed in [3] that focuses on traffic planning and safety and provide high accuracy outputs and detects anomalous traffic patterns to suggest preemptive safety measures and reduce traffic accidents and deaths. Interactive augmented reality applications must rely on object tracking, face detection, and other video analytics to obtain spatial knowledge, and must rely on cloudlet based edge solution to provide users with a smooth interaction experience[3].

Scientists at MIT's Computer Science and Artificial Intelligence Lab (CSAIL) are working on self-folding printed robots and their use in saving lives as an alternative to invasive surgery procedures, which would require a cloud in the proximity as those robots and sensors generate a large amount of data that needs to be processed very fast[4].

Verizon created a universal cloud-in-a-box solution running Linux on a generic x86 architecture, in an OpenStack container that can put compute, storage and networking resources near the edge to support their increasing number of users and power 5G in the future[4][9].

4   RECENT DEVELOPMENTS

The need for a holistic data analytics platform to combine various techniques in cloud and edge computing and to ease data management arises in applications like health monitoring where anomalies in a patient's conditions must be immediately reported and an analysis on historical patient data is needed to find out more details about patient's overall condition[10]. The lack of platforms for the edge that allow development and simple deployment in a distributed setting is a key limitation to using edge devices effectively[13].

Taking on these challenges a serverless platform was proposed by Nastic et al that supports real-time analytics across cloud and the edge by optimizing the placement of analytics operations and automatically managing available resources[10].The model takes a top-down approach for control processes combined with a bottom-up approach for data management allowing analytics to be done at various levels of granularity and the results served to the application from either the edge or the cloud depending on the need[10]. The platform provides developers with an API that allows them to easily define analytics functions without worrying about data management and optimization complexities.

Early IoT infrastructures were heavily cloud dependent and all the computation was done in the cloud. This tight coupling with the cloud is however not desirable in many time-critical or data-intensive applications[5]. An edge offloading architecture named FADES (Function virtualizAtion basED System)was proposed by Cozzolino et al that reverses the traditional paradigm and dispatches some computation to the devices close to the edge. How this offloading to the edge should be performed depends on the application, the hardware capabilities, and the software requirements[5]. The multilayer pipeline ensures reduced amount of data to be uploaded

and the MirageOS based unikernel approach provides an additional layer of security by running the deployed tasks inside a virtualization platform, bridging the gap between complex cloud-based applications and edge applications and providing modularity[5].

5   AI ON THE EDGE

With the emergence of decentralized applications, smart machines that rely on machine learning and mesh computing to provide local real-time analytics are becoming a reality. MIT's Eyeriss which is an accelerator for deep neural networks uses no wifi and no data transmission. With peer to peer networks gaining importance, edge computing is vital to provide low latency applications that are decentralized[2].

Since many artificial intelligence (AI) applications need a huge amount of processing power and require a large amount of data, traditional AI applications rely on cloud servers to perform their computation. This is a serious limitation in applications where connectivity is not reliable and time-critical decisions are required[6]. iEx.ec is a company that uses Etherium blockchain to create a market for computing resources, in turn, facilitating distributed machine learning[8].

In applications like flying a swarm of drones, a loss of connectivity to the cloud can be fatal and cause disruption of the operation. Thus AI coprocessor chips that can run machine learning algorithms can offer intelligence at the edge devices. Movidius recently announced a deep learning compute stick[1] that can add machine learning capabilities to computers and raspberry pis as a plug and play device[6].

Machine learning algorithms line one-shot learning which require lesser data are rapidly enabling edge devices to perform intelligent tasks easily[15]. Gamalon, backed by Defense Advanced Research Projects Agency (DARPA), is using Bayesian Program Synthesis to reduce the amount of data required for machine learning[6].

6   CONCLUSION

With the increase in the number of connected devices and the increase in the demand of real-time and interactive applications, we see that edge computing is a necessity and many industries are rapidly moving towards edge solutions. Although industrial IoT still faces challenges with the integration of legacy applications and proprietary applications with new technology, open source solutions are being widely accepted. Research on various platforms and architectures for edge computing continuously aims to reduce the gap between cloud and edge devices and establish standards for the same. The emergence of decentralized applications and the growing importance of machine learning has driven technologies that provide machine learning capabilities to edge devices which are becoming a fundamental requirement to move towards decentralized AI applications, that can provide results in near real time.

The author would also like to thank other associate instructors of the class for helping promptly with queries on piazza which helped everyone a great deal.

REFERENCES

[1] [n. d.]. movidius website. ([n. d.]). https://developer.movidius.com/
[2] Scott Amyx. 2016. Ready for the disruption from edge computing? IBM iot blog. (August 2016). https://www.ibm.com/blogs/internet-of-things/edge-computing/
[3] G. Ananthanarayanan, P. Bahl, P. Bodk, K. Chintalapudi, M. Philipose, L. Ravindranath, and S. Sinha. 2017. Real-Time Video Analytics: The Killer App for Edge Computing. *Computer* 50, 10 (2017), 58–67. https://doi.org/10.1109/MC.2017.3641638
[4] Jason Baker. 2017. Why OpenStack is living on the edge. opensource.com blog. (May 2017). https://opensource.com/article/17/5/openstack-summit-news
[5] Vittorio Cozzolino, Aaron Yi Ding, and Jörg Ott. 2017. FADES: Fine-Grained Edge Offloading with Unikernels. In *Proceedings of the Workshop on Hot Topics in Container Networking and Networked Systems (HotConNet '17)*. ACM, New York, NY, USA, 36–41. https://doi.org/10.1145/3094405.3094412
[6] Ben Dickson. 2017. How do you bring artificial intelligence from the cloud to the edge? TNW website. (August 2017). https://thenextweb.com/contributors/2017/08/21/bring-artificial-intelligence-cloud-edge/#.tnw_5VcrJGrz
[7] Andrew Foster. 2017. Why the Industrial IoT Needs an Open-Source Edge Platform. (july 2017). https://www.rtinsights.com/why-the-industrial-iot-needs-an-open-source-edge-platform/
[8] iEx.ec. 2017. *Building a Fully Distributed Cloud for Blockchain based Distributed Applications*. white paper. iEx.ec. http://iex.ec/wp-content/uploads/2017/04/iExec-WPv2.0-English.pdf
[9] Nicole Martinelli. 2017. Pushing the edges with OpenStack. Open Stack Articles. (May 2017). http://superuser.openstack.org/articles/edge-computing-verizon-openstack/
[10] S. Nastic, T. Rausch, O. Scekic, S. Dustdar, M. Gusev, B. Koteska, M. Kostoska, B. Jakimovski, S. Ristov, and R. Prodan. 2017. A Serverless Real-Time Data Analytics Platform for Edge Computing. *IEEE Internet Computing* 21, 4 (2017), 64–71. https://doi.org/10.1109/MIC.2017.2911430
[11] Avi Patwardhan. 2016. Incorporate streaming analytics in the Internet of Things. IBM data and analytics hub blog. (october 2016). http://www.ibmbigdatahub.com/blog/incorporate-streaming-analytics-internet-things
[12] Mahadev Satyanarayanan. 2017. The Emergence of Edge Computing. *Computer, IEEE compter society* (2017). http://elijah.cs.cmu.edu/DOCS/satya-edge2016.pdf
[13] Yogesh Simmhan. 2017. IoT Analytics Across Edge and Cloud Platforms. IEEE IOT Newsletter. (May 2017). https://iot.ieee.org/newsletter/may-2017/iot-analytics-across-edge-and-cloud-platforms.html
[14] Wikipedia. 2017. Edge computing — Wikipedia, The Free Encyclopedia. (2017). https://en.wikipedia.org/w/index.php?title=Edge_computing&oldid=802381553 [Online; accessed 7-October-2017].
[15] Wikipedia. 2017. One-shot learning — Wikipedia, The Free Encyclopedia. (2017). https://en.wikipedia.org/w/index.php?title=One-shot_learning&oldid=793877024 [Online; accessed 7-October-2017].
[16] Yang Zhao, Jeff Ashe, David Toledano, Brandon Good, Li Zhang, and Adam McCann. 2016. Occupancy and Activity Monitoring with Doppler Sensing and Edge Analytics: Demo Abstract. In *Proceedings of the 14th ACM Conference on Embedded Network Sensor Systems CD-ROM (SenSys '16)*. ACM, New York, NY, USA, 322–323. https://doi.org/10.1145/2994551.2996543

7 BIBTEX ISSUES

Warning–no key, author in movidius

Warning–no author, editor, organization, or key in movidius

Warning–to sort, need author or key in movidius

Warning–no key, author in movidius

Warning–no key, author in movidius

Warning–no key, author in movidius

Warning–no author, editor, organization, or key in movidius

Warning–empty author in movidius

Warning–empty year in movidius

Warning–numpages field, but no articleno or eid field, in FADES-offloading

Warning–no number and no volume in satyanarayananemergence

Warning–page numbers missing in both pages and numpages fields in satyanarayananemergence

Warning–numpages field, but no articleno or eid field, in Dolppler-usecase

(There were 13 warnings)

8 ISSUES

DONE:

Example of done item: Once you fix an item, change TODO to DONE

8.1 Formatting

Incorrect number of keywords or HID and i523 not included in the keywords

8.2 Citation Issues and Plagiarism

Put a space between the citation mark and the previous word

Are you sure you have quoted all the directly cited material